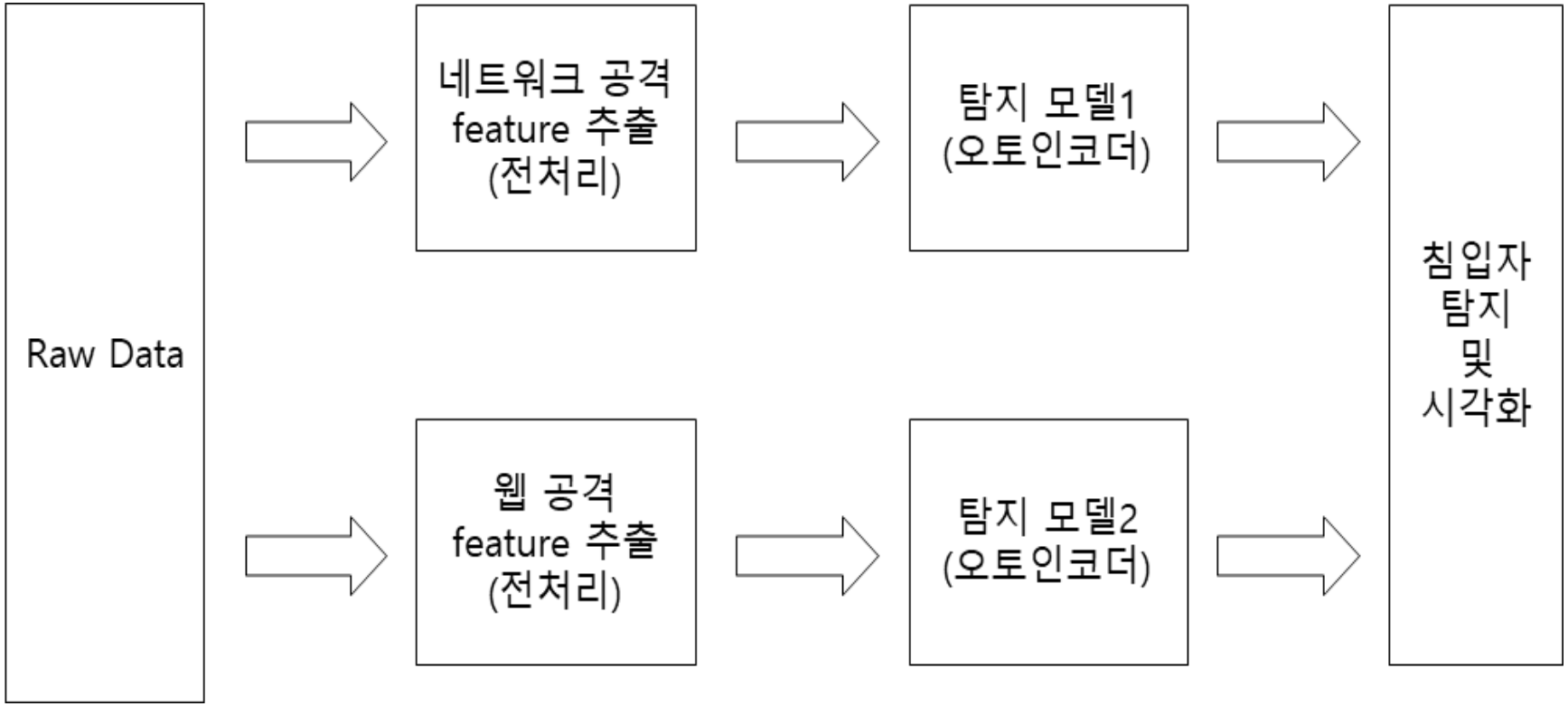
1. 시스템 구성도


Raw Data
네트워크 공격
feature 추출
(전처리)
웹 공격
feature 추출
(전처리)
탐지 모델1
(오토인코더)
탐지 모델2
(오토인코더)
침입자
탐지
및
시각화

2. 네트워크 공격 탐지 모델

Port Scan, DOS, Brute Force 등 탐지

유일 키

- 출발지IP, 도착지IP, 출발지 Port, 도착지 Port
- 세션 시작 시간, 세션 종료 시간

파생 변수1 - 기본 Feature에서 추출

- 방향, 출발지 포트가 알려진 포트, 도착지 포트가 알려진 포트
- 세션 유지 시간
- 출발지에서 보낸 패킷의 개수, 도착지에서 보낸 패킷의 개수
- 출발지에서 보낸 패킷의 크기, 도착지에서 보낸 패킷의 크기
- 프로토콜, 서버 응답 값

파생 변수2 - Time series 속성

- 포트 스캔, IP스캔
- 전/후 60개 세션 중 출발지에서 보낸 패킷의 개수 합
- 전/후 60개 세션 중 도착지에서 보낸 패킷의 개수 합
- 전/후 60개 세션 중 출발지에서 보낸 패킷의 크기 합
- 전/후 60개 세션 중 도착지에서 보낸 패킷의 크기 합
- 전/후 60개 세션 중 응답 값 카운트 합

3. 웹 공격 탐지 모델

SQL Injection, XSS, bot 등 탐지

정적 변수

- **user-agent, content-type**
- **데이터 기반으로 단어장 생성하여 단어장에 없으면 탐지**

정적 & 동적 변수

- **http request url**
- **정적인 부분은 디렉토리 경로**
- **동적인 부분은 url에서 매개변수로 전달하는 값**

3. 웹 공격 탐지 모델

SQL Injection, XSS, bot 등 탐지

빨강: 정적 변수
파랑: 동적 변수

/abc/def?q=123&id=abcd

동적 파생변수 추출 특성

- 글자 수, 특수문자 수(공격에 자주 사용된 특수문자 예를 들어, ‘ # @ 등)
  영어 소문자 수, 영어 대문자 수, 숫자 수, 띄어쓰기 수(+, \t 등)
  호스트 서버 주소, 매개변수 이름
- 일반적으로 웹 공격은 매개 변수에 공격 코드를 추가하여 발생하기 때문에
  평소와는 형태가 많이 다른 매개변수가 입력되는 경우 공격으로 판단

3. 웹 공격 탐지 모델

SQL Injection, XSS, bot 등 탐지

정상

/dv/vulnerabilities/xss_r/?name=TEST

비정상

/dv/vulnerabilities/xss_r/?name=<script>console.log('01CGBIIDYQA0C94JIW0U8A08G01F5FVYRQ5L2YJSMYPMPQB34N');console.log(document.cookie);</script>

글자 수, 특수문자 수(공격에 자주 사용된 특수문자, 예를 들어, '#@ 등), 영어 소문자 수, 영어 대문자 수, 숫자 수, 띄어쓰기 수(+, \t 등)을 특성으로 추출하여 이상 값 탐지

4. 평가

o F1-score 수식으로 탐지 정확도 측정

$$F1\text{-}score = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

o Precision (정확률) : 제안한 알고리즘이 공격을 판단했을 때 실제 공격이 이루어진 비율 $\left(\frac{\text{정답 인정수}}{\text{참가자가 예측한 공격 수}}\right)$

o Recall (재현율) : 실제 공격이 이루어졌을 때 제안된 알고리즘이 공격을 판단하는 비율 $\left(\frac{\text{정답 인정수}}{\text{실제 공격 수}}\right)$

카테고리		정답	
		공격	정상
탐지결과	공격	True Positive (TP)	False Positive (FP)
	정상	False Negative (FN)	True Negative (TN)

$$Precision = \frac{TP}{TP+FP} \qquad Recall = \frac{TP}{TP+FN}$$

4. 결과

SET1 탐지율

-> 89%

SET2 탐지율

-> 59%

실제 운영환경에서는 더 높은 탐지율을 기록할 것으로 예상함.(실제 환경에서는 웹 공격이 다수이므로)

웹 공격 탐지 모델은 **실시간 탐지** 가능하며 네트워크 공격 탐지 모델은 **배치** 형태로 탐지 가능.

방화벽 로그, 웹 서버 로그를 이용하여 모델 구축 가능하기 때문에 **기존의 보안 시스템과 조화로운 운영** 가능.